EVOLUCIÓN DEL PNB AMERICANO DATOS TRIMESTRALES 1954:Q1-1987:Q4

En este trabajo práctico, nos interesamos en la serie del PNB estadounidense (su logaritmo se indica con más precisión (X_t) a continuación) durante el período 1954-1987 en dólares. Hacemos

X_t = PNB=log(GNP)

Comience cargando el paquete "tseries" y luego

```
library(tseries)
data(USeconomic)
PNB = USeconomic[, 2]
annee = seq(1954, 1987.75, 0.25)
plot(annee, PNB, main = "log(PNB)a lo largo del tiempo", t = "l", col = "blue", xlab = "tiempo", ylab
= "log PNB")
```

1 Estacionariedad de la serie log (PNB)

Tracemos el correlograma y el correlograma parcial.

```
par(mfrow = c(1, 2))
acf(PNB)
pacf(PNB)
```

El análisis de la función de autocorrelación confirma la no estacionariedad de la serie log(PNB): de hecho, observamos una persistencia y la función de autocorrelación empírica ρ(h) no es cero para todos los términos h ≥ 0. La prueba de maleta (que prueba la no correlación de la muestra) también rechaza la estacionariedad de la serie:

```
Box.test(PNB)
```

2 Estudio de la serie diff(log(PNB))

Diferenciación, estacionariedad y elección de modelo.

Ahora trabajaremos sobre la serie $(Y_t = X_t - X_{t-1})_{2 \leq t \leq T}$. En el trazo de Y_t, vemos que la serie ya no parece tener tendencia (comportamiento muy errático).

```
diffPNB = diff(PNB)
T = length(annee)
par(mfrow = c(1, 1))
plot(annee[2 :T], diffPNB, main = "diffPNB", t = "l",col = "blue",xlab = "tiempo", ylab = "diff PNB")
```

También podemos trazar las funciones de autocorrelación empíricas.

```
par(mfrow = c(1, 2))
acf(diffPNB, lag.max = 30)
pacf(diffPNB, lag.max = 30)
```

Como las funciones de autocorrelación (normal y parcial) disminuyen muy rápidamente, podemos considerar que la serie de diferencias Y_t es estacionaria. Además, la serie Y_t no está centrada a pesar de una media empírica baja de 0.007596586. Esto también lo podemos ver contrastando la hipótesis $"\overline{Y}_T = 0"$ con el test de Student: rechazamos la hipótesis nula:

mean(diffPNB)
t.test(diffPNB)

Por tanto, necesitamos un término constante en el modelado de Y. También podríamos aplicar el operador Δ nuevamente para eliminar este término promedio y centrar la serie.

Dado que Y_t es estacionario, esto significa que la serie logarítmica del PNB X_t es del tipo paseo aleatorio ya que $Y_t = X_t - X_{t-1}$ es estacionaria. A la vista de los autocorrelogramas, un modelo ARMA(0,2) parece ser el más adecuado. También podemos elegir los modelos ARMA(1,2) o ARMA(8,2) o incluso ARMA(0,1). De hecho, la función de autocorrelación ρ(h) está fuera del intervalo de confianza para h = 1 y h = 2 y la función de autocorrelación parcial τ(h) está fuera del intervalo de confianza para h = 1 y casi fuera del intervalo para h = 8.

3 Estimación de modelos

La estimación de máxima verosimilitud se realiza con la función arima de R.

arima(diffPNB, c(0, 0, 2))
Estimemos también los parámetros de los otros tres modelos elegidos:

arima(diffPNB, c(1, 0, 2))
arima(diffPNB, c(8, 0, 2))
arima(diffPNB, c(0, 0, 1))

4 Estacionariedad de los residuos

Para comparar los modelos entre sí, buscamos minimizar el criterio AIC (Criterio de Información de Aikake). El modelo con el criterio AIC más bajo (-863,73) es el modelo ARMA(0,2). Por tanto, parece ser el que mejor se ajusta a los datos. Sin embargo, también podemos comparar estos 4 modelos probando la estacionariedad de sus residuos.

```
residus02 = arima(diffPNB, c(0, 0, 2))$residual
Box.test(residus02)
residus12 = arima(diffPNB, c(1, 0, 2))$residual
Box.test(residus12)
residus82 = arima(diffPNB, c(8, 0, 2))$residual
Box.test(residus82)
residus01 = arima(diffPNB, c(0, 0, 1))$residual
Box.test(residus01)
```

Vemos que el modelo ARMA(1,2) parece mejor, ya que el valor p de la prueba de Box Pierce es mayor. Por tanto, no podemos rechazar el hecho de que los residuos no estén correlacionados. La gráfica de los autocorrelogramas confirma que los modelos ARMA(1,2), ARMA(8,2) y ARMA(0,2) se ajustan bien a los datos.

```
par(mfrow = c(4, 2))
acf(residus02, main = "modelo ARMA(0,2)")
pacf(residus02, main = "modelo ARMA(0,2)")
acf(residus12, main = "modelo ARMA(1,2)")
pacf(residus12, main = "modelo ARMA(1,2)")
acf(residus82, main = "modelo ARMA(8,2)")
pacf(residus82, main = "modelo ARMA(8,2)")
acf(residus01, main = "modelo ARMA(0,1)")
pacf(residus01, main = "modelo ARMA(0,1)")
```

6 Normalidad de los residuos

Ahora probamos la normalidad de los residuos.

```
shapiro.test(residus02)
shapiro.test(residus12)
shapiro.test(residus82)
shapiro.test(residus01)
```

Sólo el modelo ARMA(0,1) rechaza la normalidad de los residuos al 95%.

7 Pronóstico

Ahora probaremos el rendimiento de los 4 modelos. Para ello, eliminamos los últimos 10 puntos de nuestra serie Y_t, para luego comparar nuestros 10 valores predichos con nuestros 10 valores reales.

```
n = 10
index = 1 :(T - n - 1)
pred02 = predict(arima(diffPNB[index], c(0, 0, 2)), n)

plot(annee[(T - 4 * n) :T], diffPNB[(T - 4 * n) :T - 1], main ="prevision ARMA(0,2)", t = "l", col =
"blue", xlab = "temps", ylab = "diff PNB")
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred02$pred))
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred02$pred) + c(0,pred02$se)* 1.96, lty = 2)
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred02$pred) - c(0,pred02$se)* 1.96, lty = 2)

n = 10
index = 1 :(T - n - 1)
pred03 = predict(arima(diffPNB[index], c(1, 0, 2)), n)

plot(annee[(T - 4 * n) :T], diffPNB[(T - 4 * n) :T - 1], main ="prevision ARMA(1,2)", t = "l", col =
"blue", xlab = "temps", ylab = "diff PNB")
```

```
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred03$pred))
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred03$pred) + c(0,pred02$se)* 1.96, lty = 2)
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred03$pred) - c(0,pred02$se)* 1.96, lty = 2)

n = 10
index = 1 :(T - n - 1)
pred04 = predict(arima(diffPNB[index], c(8, 0, 2)), n)

plot(annee[(T - 4 * n) :T], diffPNB[(T - 4 * n) :T - 1], main ="prevision ARMA(8,2)", t = "l", col =
"blue", xlab = "temps", ylab = "diff PNB")
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred04$pred))
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred04$pred) + c(0,pred02$se)* 1.96, lty = 2)
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred04$pred) - c(0,pred02$se)* 1.96, lty = 2)

n = 10
index = 1 :(T - n - 1)
pred05 = predict(arima(diffPNB[index], c(0, 0, 1)), n)

plot(annee[(T - 4 * n) :T], diffPNB[(T - 4 * n) :T - 1], main ="prevision ARMA(0,1)", t = "l", col =
"blue", xlab = "temps", ylab = "diff PNB")
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred05$pred))
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred05$pred) + c(0,pred02$se)* 1.96, lty = 2)
lines(annee[(T - n) :T], c(diffPNB[T - n - 1], pred05$pred) - c(0,pred02$se)* 1.96, lty = 2)
```

Haz lo mismo con los otros 3 modelos. Observando los gráficos, deducimos que las predicciones del modelo ARMA(0,1) son peores que las de los otros dos modelos. Los intervalos de confianza (95%) de los valores predichos suponen que los residuos son gaussianos (lo que parece ser cierto para los modelos ARMA(1,2), ARMA(8,2) y ARMA (0,2), ya que no se puede rechazar la hipótesis de normalidad).

Como conclusión del análisis de la serie diferenciada Y_t, nos llevamos a preferir el modelo ARMA(0,2), porque minimiza el criterio AIC sin perder la estacionariedad de los residuos y requiere menos parámetros.

8 Modelo ARIMA para la serie log(PNB)

Diferenciación de segundo orden y elección de modelo.

La serie Y_t de diferencias en la serie logarítmica del PNB nos parecía centrada, pero si dibujamos el histograma vemos que no es nada simétrico.

```
hist(diffPNB)
```

Así, diferenciaremos la serie logarítmica del PNB por segunda vez: fijamos

$Z_t = \Delta^2 X_t (= \Delta Y_t)$. Vemos que la serie es más centrada y sobre todo más simétrica.

```
d2PNB = diff(diffPNB)
par(mfrow = c(1, 2))
plot(annee[3 :T],d2PNB,main="Z a lo largo del tiempo",t="l",col="blue",xlab="temps",ylab="Z")
hist(d2PNB)

mean(d2PNB)
```

La prueba de Student revela que no podemos rechazar la hipótesis de que $\overline{Z}_T = 0$.

Podríamos estudiar directamente la serie d2PNB como se explicó anteriormente, pero luego tendríamos que volver a PNB para determinar sus predicciones. Por lo tanto, elegimos aquí trabajar directamente en PNB y modelar esta serie utilizando un ARIMA. En cuanto a los modelos ARMA, ahora encontraremos los coeficientes p y q del modelo y luego probaremos la estacionariedad de los residuos y el rendimiento de nuestras predicciones.

```
par(mfrow = c(1, 2))
acf(d2PNB, lag.max = 30)
pacf(d2PNB, lag.max = 30)
```

Trazar los autocorrelogramas revela que q = 1 y p = 5 o 9.

9 Estimación de modelos

Por lo tanto, probaremos dos modelos ARIMA(5,2,1) y ARIMA(9,2,1) o posiblemente si queremos ser económicos con los parámetros ARIMA(1,2,1). Estimamos los coeficientes:

```
arima(PNB, c(1, 2, 1), inc = FALSE)
arima(PNB, c(5, 2, 1), inc = FALSE)
arima(PNB, c(9, 2, 1), inc = FALSE)
```

Según el criterio AIC, el modelo ARIMA(1,2,1) parece ajustarse mejor a los datos que el modelo ARIMA(5,2,1).

10 Estacionariedad y normalidad de los residuos.

Probemos la estacionariedad y normalidad de los residuos.

```
residus121=arima(PNB, c(1, 2, 1), inc = FALSE)$residual
residus521=arima(PNB, c(5, 2, 1), inc = FALSE)$residual
residus921=arima(PNB, c(9, 2, 1), inc = FALSE)$residual
Box.test(residus121)
Box.test(residus521)
Box.test(residus921)
shapiro.test(residus121)
shapiro.test(residus521)
shapiro.test(residus921)
```

De esto deducimos que no podemos rechazar la estacionariedad y normalidad de los residuos para los tres modelos considerados. Sin embargo, para el modelo ARIMA(1,2,1) el valor p sigue siendo bajo. Por lo tanto, preferiremos los modelos ARIMA(5,2,1) y ARIMA(9,2,1) que ciertamente tienen más parámetros pero buen AIC y residuos estacionarios (valor p muy cercano a 1). La hipótesis de normalidad aún no es muy probable (valores p bajos de Shapiro). Finalmente conservaremos el modelo ARIMA(5,2,1) que tiene el AIC más pequeño y la menor cantidad de parámetros.

```
acf(residus121)
acf(residus521)
acf(residus921)
pacf(residus121)
pacf(residus521)
pacf(residus921)
```

Finalmente, la gráfica de las funciones de autocorrelograma no muestra ninguna correlación para los residuos. Al final, el modelo ARIMA(5,2,1) parece mejor que ARIMA(9,2,1) según estos datos.

11 Pronóstico

Finalmente, probemos el desempeño de las predicciones.

```
index = 1 :(T - n)
pred521 = predict(arima(PNB[index], c(5, 2, 1), inc = FALSE),n)
pred921 = predict(arima(PNB[index], c(9, 2, 1), inc = FALSE),n)
plot(annee[(T - 4 * n) :T], PNB[(T - 4 * n) :T], main = "prevision ARIMA(5,2,1)",t = "l", col = "blue",
xlab = "temps", ylab = "PNB")
lines(annee[(T - n) :T], c(PNB[T - n], pred521$pred))
lines(annee[(T - n) :T], c(PNB[T - n], pred521$pred) + c(0, pred521$se)* 1.96, lty = 2)
lines(annee[(T - n) :T], c(PNB[T - n], pred521$pred) - c(0, pred521$se)* 1.96, lty = 2)
```

Repite lo mismo para el otro modelo. Considerando los pronósticos, llegamos a la conclusión de que el modelo ARIMA(5,2,1) se adapta mejor a los datos.

12 Conclusión

Las siguientes tablas resumen todos los resultados de los diferentes modelos de TP.

Figura 2 – Resumen: criterio AIC, probabilidad logarítmica y pruebas de residuos para ARMA

Critères-Modèles	ARMA(0,2)	ARMA(0,1)	ARMA(1,2)	ARMA(8,2)
AIC	-863,73	-860,05	-861,89	-855,32
Vraisemblance	435,86	433,03	439,62	439,66
p-valeur Box Pierce	0,9474	0,5948	1	0,9997
p-valeur Shapiro	0,323	0,2492	0,4097	0,1872

Figura 3 – Resumen: criterio AIC, probabilidad logarítmica y pruebas de residuos para ARIMA

Critères-Modèles	ARIMA(1,2,1)	ARIMA(5,2,1)	ARIMA(9,2,1)
AIC	-851,52	-847,31	-843,02
Vraisemblance	428,76	430,66	432,51
p-valeur Box Pierce	0,6096	0,9401	0,9561
p-valeur Shapiro	0,1175	0,2780	0,2963

En conclusión, el modelo ARMA(0,2) en la serie PNB diferenciada es mejor dados los criterios elegidos que el modelo ARIMA(5,2,1) en la serie PNB. Sin embargo, si optamos por trabajar con la serie PNB diferenciada y por tanto con el modelo ARMA(0,2), el software R nos proporcionará previsiones para la serie diferenciada. Entonces depende de nosotros calcular manualmente los predictores de la serie PNB inicial escribiendo que diffPNB = (I − B)PNB. Por tanto, es mejor trabajar directamente con el modelado ARIMA.